

## Morfometria do trato vocal por ressonância magnética: simulação de padrões patológicos articulatórios

Sandra M. Rua Ventura<sup>1</sup>, Diamantino Rui S. Freitas<sup>2</sup>, Isabel Maria A. P. Ramos<sup>3</sup>, João Manuel R. S. Tavares<sup>4</sup>

1. Área Técnico-Científica da Radiologia, Escola Superior de Tecnologia da Saúde do Porto, Instituto Politécnico do Porto.

2. Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia, Universidade do Porto.

3. Serviço de Radiologia do Centro Hospitalar de S. João, EPE, Faculdade de Medicina, Universidade do Porto.

4. Instituto de Engenharia Mecânica e Gestão Industrial, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto. tavares@fe.up.pt

**RESUMO:** **Introdução** – A análise da forma ou morfometria de estruturas anatómicas, como o trato vocal, pode ser efetuada a partir de imagens bidimensionais (2D) como de aquisições volumétricas (3D) de ressonância magnética (RM). Esta técnica de imagem tem vindo a ter uma utilização crescente no estudo da produção da fala. **Objetivos** – Demonstrar como pode ser efetuada a morfometria do trato vocal a partir da imagem por ressonância magnética e ainda apresentar padrões anatómicos normais durante a produção das vogais [i a u] e dois padrões articulatórios patológicos em contexto simulado. **Métodos** – As imagens consideradas foram recolhidas a partir de aquisições 2D (*Turbo Spin-eco*) e 3D (*Flash Gradiente-Eco*) de RM em quatro sujeitos durante a produção das vogais em estudo; adicionalmente procedeu-se à avaliação de duas perturbações articulatórias usando o mesmo protocolo de RM. A morfometria do trato vocal foi extraída com recurso a técnicas manuais (para extração de cinco medidas articulatórias) e automáticas (para determinação de volumes) de processamento e análise de imagem. **Resultados** – Foi possível analisar todo o trato vocal, incluindo a posição e a forma dos articuladores, tendo por base cinco medidas descritivas do posicionamento destes órgãos durante a produção das vogais. A determinação destas medições permitiu identificar quais as estratégias mais comumente adotadas na produção de cada som, nomeadamente a postura articulatória e a variação de cada medida para cada um dos sujeitos em estudo. No contexto de voz falada intersujeitos, foi notória a variabilidade nos volumes estimados do trato vocal para cada som e, em especial, o aumento do volume do trato vocal na perturbação articulatória de sigmatismo. **Conclusão** – A imagem por RM é, sem dúvida, uma técnica promissora no estudo da fala, inócua, não-invasiva e que fornece informação fiável da morfometria do trato vocal.

*Palavras-chave: imagiologia do trato vocal, imagem por ressonância magnética (IRM), análise morfométrica, articulação da fala, processamento e análise de imagem, perturbações articulatórias.*

## Vocal tract morphometry by magnetic resonance imaging: simulation of pathological articulatory patterns

**ABSTRACT:** **Introduction** – The shape or morphologic analysis of anatomical structures, such as the vocal tract can be performed from two-dimensional (2D) or volumetric acquisitions (3D) of magnetic resonance imaging (MRI). This imaging technique has had an increasing use in the study of speech production. **Objectives** – To determine a method to perform the morphometric analysis of the vocal tract from magnetic resonance imaging; to present anatomical patterns during the normal speech production of some vowels and two pathological articulatory disorders in simulated context. **Methods** – The image data was collected from 2D (*Turbo Spin Echo*) and 3D (*Flash Gradient Echo*) acquisitions of MRI of four subjects during the production of three vowels; in addition, two articulatory disorders were assessed using this imaging proto-

col. The morphology of the vocal tract was extracted using manual and automatic techniques of image processing and analysis. **Results** – Based on five articulatory measurements, it was possible to study the entire vocal tract during vowel production, including the position and shape of the articulators involved. Based on these measurements, it was possible to identify the strategies most commonly adopted in the production of each sound, including the articulatory posture and the modification of each measure for the subjects under study. Concerning the voices of the different speakers, the variability in the assessed volumes of the vocal tract for each sound was found, and in particular, the increased vocal tract volume in the articulatory disorder – the sigmatism. **Conclusion** – MRI is a promising technique for speech production studies, safe, non-invasive and that provides reliable information concerning the morphometric analysis of the vocal tract.

*Keywords: vocal tract imaging, magnetic resonance imaging (MRI), morphometric analysis, speech articulation, image processing and analysis, articulatory disorders.*

## Introdução

O interesse pelo conhecimento da produção da fala surge paralelamente com o desenvolvimento de novas modalidades imagiológicas, como a ressonância magnética (RM), nos finais do século XX, tendo o primeiro estudo<sup>1</sup> sido realizado em 1986. A observação estática dos articuladores por RM durante a sustentação muito prolongada dos sons rapidamente foi aperfeiçoada, permitindo não só uma maior extensão de sons estudados, mas também o alargamento a várias línguas, como o francês<sup>2-3</sup>, inglês<sup>4-5</sup>, alemão<sup>6</sup>, japonês<sup>7</sup> e o sueco<sup>8-9</sup>.

Impôs-se também a necessidade de incrementar o rigor desta avaliação a partir de imagens bidimensionais (2D) e mais tarde, tridimensionais (3D), quer para caracterização morfológica como dinâmica<sup>4</sup> da produção da fala.

Por oposição aos estudos da área da Engenharia<sup>10-14</sup>, usualmente com um número reduzido de participantes privilegiando um *corpus*\* da fala mais vasto, os estudos na vertente clínica<sup>15-17</sup> são menos frequentes, incluem mais participantes e assentam quase exclusivamente na caracterização funcional da fala.

Pela revisão da literatura é notório o interesse que os estudos da produção da fala despertam em várias áreas multidisciplinares<sup>18-19</sup>, assim como pela explicação de outros fenómenos como a nasalidade ou a coarticulação<sup>20-21</sup>. De realçar também a evolução e aperfeiçoamento da técnica de RM e o desenvolvimento de diferentes abordagens na procura de um melhor entendimento da fala<sup>21</sup>.

Apesar dos vários estudos existentes sobre a produção da fala, a informação morfológica e dinâmica é ainda exígua, nomeadamente no que diz respeito aos sons do português europeu (PE)<sup>20,22</sup>. Assim, a primeira utilização da RM<sup>19</sup> resultou na caracterização morfológica de 25

sons do PE e na recolha de cine-imagens do trato vocal durante a fala em contexto articulatorio consoante-vogal (CV) para as sílabas /pa/, /ma/ e /tu/, impulsivando a expansão e aperfeiçoamento da técnica de RM por outros estudos.

Nos anos seguintes, um outro estudo apresentou uma aquisição ultrarrápida de RM<sup>20</sup>, conseguindo dados volumétricos de um sujeito para estudo das vogais orais e nasais e das consoantes fricativas, laterais e oclusivas. Neste estudo, as imagens bidimensionais reconstruídas a partir das aquisições volumétricas permitiram o estudo comparativo entre as vogais orais e nasais, mediante a extração de contornos e a posterior correlação com uma base de dados de articulografia eletromagnética desenvolvida anteriormente. O estudo tridimensional realizado permitiu a determinação das funções de área de parte deste *corpus*, tendo-se mostrado particularmente útil na caracterização das consoantes palatais.

A aplicação da RM e potencialidades no estudo do PE foram descritas num outro estudo<sup>22</sup>, recorrendo à aquisição de imagens em diferentes orientações. Este culminou na apresentação de uma base de dados de natureza morfológica do trato vocal em várias posições articulatorias e para um número considerável de sons do PE, constituindo os primeiros modelos 3D do trato vocal reconstruídos a partir de imagens de RM recolhidos em dois sujeitos jovens, voluntários e saudáveis. Este estudo demonstrou o valor da técnica de RM não só em medicina, contribuindo para o estudo de disfunções articulatorias e no planeamento terapêutico da fala, como em engenharia para o reconhecimento e síntese da fala.

Um estudo realizado para caracterização das consoantes laterais<sup>23</sup>, em diferentes contextos silábicos, descreveu as características articulatorias mais comumente encontradas, apontando como principal limitação do estudo a dificuldade de registo espacial e segmentação dos modelos de gesso dos dentes.

\* O *corpus* da fala consiste num conjunto de sons da fala recolhidos criteriosamente para serem objeto de pesquisa.

Em 2011, outro estudo apresentou a descrição de um estudo dinâmico por RM<sup>24</sup>, realizado em dois sujeitos, um do sexo masculino e um outro do sexo feminino, usando um *pseudo-trigger* eletrocardiográfico para sincronização de imagens cinemáticas de RM. Neste estudo foram calculados quatro parâmetros articulatórios usados frequentemente nos sistemas de síntese da fala, demonstrando-se diferenças importantes na forma do trato vocal, comparativamente entre os dois sujeitos durante a produção dos mesmos sons. Mediante a utilização de modelos estatísticos deformáveis, em particular, modelos de distribuição pontual e modelos ativos, em imagens de RM, outro estudo demonstrou a possibilidade de extração automática da forma do trato vocal para diferentes sons do PE<sup>25</sup>. Os autores concluíram que estes modelos permitem a extração de forma eficaz e automática das principais características do trato vocal durante os movimentos articulatórios.

As especificidades das imagens de RM aliadas à morfologia particular do trato vocal criam importantes condicionantes às tarefas computacionais de processamento e análise de imagem<sup>26</sup>. Adicionalmente, a área do processamento e análise de imagem carece de *softwares* específicos que permitam responder às necessidades dos estudos da fala; daí que a maioria dos autores recorra a métodos manuais para as tarefas de segmentação das imagens, isto é, para a extração dos contornos do trato vocal em estudo. Contudo, a utilização de técnicas mais automatizadas tem vindo a crescer e com resultados promissores, nomeadamente nos estudos em tempo-real<sup>25,27</sup>.

A análise da forma ou morfometria de estruturas anatómicas, como o trato vocal<sup>28</sup>, pode ser efetuada a partir de imagens bidimensionais, como também a partir de aquisições volumétricas tridimensionais, tanto de forma qualitativa, mais subjetiva, como quantitativa. Assim, com este estudo pretende-se demonstrar como pode ser efetuada a morfometria do trato vocal a partir de imagens de RM e ainda apresentar padrões anatómicos normais durante a produção das vogais [i a u] e dois padrões articulatórios patológicos em contexto simulado.

Este artigo encontra-se estruturado em quatro partes principais: descrevem-se, na secção seguinte, os participantes, o *corpus* da fala e o protocolo de RM adotado; na secção de resultados e discussão são apresentadas a morfometria bidimensional (2D) e tridimensional (3D) do trato vocal extraídas a partir das imagens de RM, exibidos modelos 3D do trato vocal durante a produção de alguns sons da fala e discutidos os resultados apresentados. Por fim, são apresentadas as conclusões finais e apontadas algumas perspetivas futuras.

## Metodologia

Este estudo foi realizado mediante a autorização e disponibilização de recursos humanos e materiais da Universidade de Ressonância Magnética do Serviço de Radiologia do Centro Hospitalar de S. João, EPE, no Porto e seguindo as normas operacionais desse serviço.

Nesta secção são descritos os materiais e os métodos utilizados, assim como as técnicas de análise morfométrica usadas.

### Participante e corpus da fala

Da amostra fizeram parte quatro profissionais da voz, dois do sexo feminino e dois do sexo masculino, saudáveis e que voluntariamente aceitaram participar, mediante consentimento informado relativamente ao estudo a realizar. Por questões de segurança e de forma a despistar eventuais contraindicações à RM, realizou-se um questionário previamente ao estudo. Dada a especificidade da técnica de RM e deste estudo em particular, procedeu-se à informação através de um guia prático elaborado para o efeito, preparação e treino prévio de cada sujeito com auxílio de um terapeuta da fala.

Assim, os participantes apresentavam idades compreendidas entre os 30 e 47 anos (média  $35,5 \pm 8,0$  anos), um peso médio de  $63,0 \pm 8,3$  Kg e sem historial de patologia oro-facial ou perturbação articulatória, segundo a avaliação realizada durante o processo de seleção por um terapeuta da fala.

Para constituição do *corpus* da fala escolheram-se as três vogais [i a u] da transcrição fonética do *International Phonetic Alphabet* (IPA) para o PE, pela familiaridade dos sons por qualquer um dos participantes e pela facilidade com que são produzidos. Acresce ainda que, uma vez que estas vogais constituem as extremidades do triângulo das vogais, ou seja, ocupando as posições mais extremas da articulação de vogais, figuram como representativas da mobilidade global do sistema. Todos os sons foram sustentados durante as aquisições de RM.

No contexto de voz falada<sup>†</sup> (voz normal de cada sujeito), as vogais foram pronunciadas utilizando uma palavra de referência em contexto articulatório CV, usando-se a consoante oclusiva [p] por não influenciar a produção dos sons vocálicos e, assim, permitir uma produção mais natural do som. As palavras de referência foram as seguintes: [papa], [pipa] e [pupa].

Para o estudo de alguns padrões patológicos articulatórios foi solicitado apenas a um dos participantes no estudo, com treino vocal e formação profissional, a simulação das seguintes perturbações: sigmatismo interdental e sigmatismo lateral, através da sustentação da sílaba /sa/ da palavra de referência /saco/ e, por fim, voz tensa para as três vogais em estudo.

† A voz falada é natural e inconsciente não necessitando de treino ou adaptações específicas, ao contrário da voz profissional.

### Equipamento utilizado e tipos de estudo

Para a aquisição das imagens utilizou-se a seguinte instrumentação: um sistema de RM Magnetom Trio (Siemens AG, Alemanha) com uma intensidade de campo de 3.0 T e duas antenas *Phased Array* para a emissão e captação do sinal da região da cabeça e pescoço, com 32 canais e 4 canais, respectivamente. A gravação da fala não foi possível dado o elevado ruído acústico produzido pelo sistema de RM durante o processo de aquisição.

Os sujeitos foram posicionados em decúbito dorsal e

com um apoio sob os membros inferiores para maior conforto.

O protocolo de RM incluiu: 1) um estudo estático realizado em todos os participantes para permitir uma caracterização morfológica 2D das cavidades de ressonância e dos articuladores; e 2) um estudo 3D orientado para a avaliação de padrões normais da fala e patológicos em contexto simulado, apenas num dos participantes do sexo masculino. A Tabela 1 resume os diferentes tipos de estudo, finalidade e parâmetros de aquisição gerais considerados.

**Tabela 1:** Protocolo de RM para estudo do trato vocal e simulação de padrões patológicos articulatorios

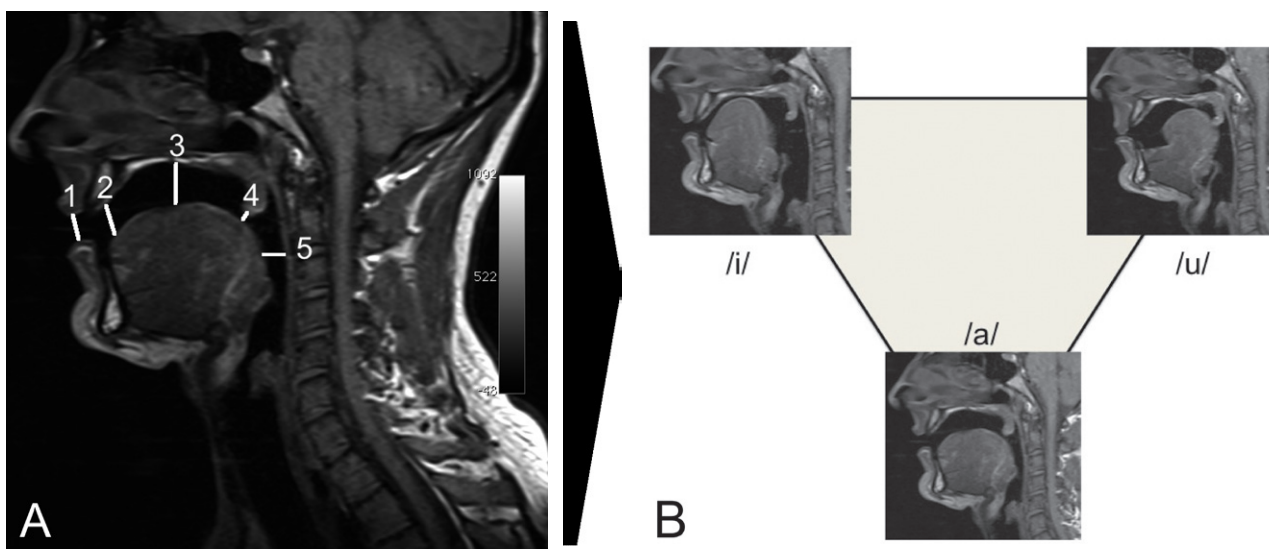
Estudo	Finalidade	Parâmetros de aquisição
2D Estático	Servir de referencial para os sons vocálicos da voz falada de cada sujeito e efetuar a morfometria 2D do trato vocal. Aplicado aos quatro participantes.	Sequência Turbo Spin-Eco (TR = 400ms, TE = 10ms, ETL = 5), um corte sagital com 3mm de espessura. FOV = 240 x 240mm, Matriz 512 x 512 e resolução de 0,469mm/pixel. Algoritmo de aquisição paralela: Não. Tempo de aquisição de 8,07s.
3D Estático	Avaliar a articulação dos sons em contexto sustentado e efetuar a morfometria 3D dos quatro participantes. Simulação dos padrões patológicos de sigmatismo e voz tensa apenas num dos sujeitos.	Sequência <i>Flash</i> Gradiente-Eco (TR = 7,6ms, TE = 2,87ms, FA = 5°). Slab no plano sagital com 78mm de espessura. FOV = 240 x 240mm, Matriz 256 x 256, resolução = 0,938mm/pixel. Total de imagens adquiridas: 26 com 3mm de espessura/ corte. Tempo de aquisição de 10,06s. Algoritmo de aquisição paralela pelo método auto-calibrativo - <i>Generalized Autocalibrating Partially Parallel Acquisition</i> (GRAPPA) Fator de aceleração – 4 Direção codificação de fase – AP (linhas)

Abreviaturas: AP – Ântero-posterior; ETL – Echo Train Length; FA- Flip Angle; FOV – Field of View; TE – Tempo de eco; TR – Tempo de repetição.

### Técnicas de análise de imagem

Para análise 2D por RM do trato vocal foram efetuadas manualmente cinco medidas posturais articulatorias: abertura labial (AL), localização do ápex da língua (LAL), localização do dorso da língua (LDL), distância da língua-véu do palato (DLV) e abertura da orofaringe (ORO)

(cf. Figura 1). Estas medidas articulatorias foram obtidas manualmente com recurso ao programa OsiriX® (Antoine Rosset, Suíça). Para a avaliação estatística da variabilidade intersujeitos realizou-se uma análise das médias e dos coeficientes de variação com recurso ao programa Microsoft® Excel® (versão 14.3.9, 2011).



1. Abertura labial (AL); 2. Localização do ápex da língua (LAL); 3. Localização do dorso da língua (LDL); 4. Distância da língua-véu do palato (DLV); 5. Abertura da orofaringe (ORO).

**Figura 1:** Medidas articulatórias efetuadas a partir do estudo estático 2D de RM (A) e imagens sagitais de RM com a respetiva representação do triângulo das vogais do PE. (B) O triângulo articulatório das vogais representa os pontos de articulação da língua com os outros articuladores para a produção desses sons: anterior e alta (vogal /i/), central (vogal /a/) e posterior e alta (vogal /u/).

A análise morfométrica 3D consistiu na extração da volumetria do trato vocal e posterior cálculo dos seus volumes a partir das aquisições volumétricas de RM com recurso ao programa ITK-SNAP® (*Penn Image Computing and Science Laboratory, Radiology Department, University of Pennsylvania*). Para o efeito realizou-se a segmentação automática das imagens e a reconstrução 3D do trato vocal a partir das mesmas.

## Resultados e discussão

De acordo com o protocolo de RM utilizado recolheram-se um total de 12 imagens sagitais para caracterização morfológica 2D articulatória dos sons em estudo e 17 aquisições volumétricas, perfazendo um total de 442 imagens sagitais para a análise morfométrica 3D do trato vocal.

As imagens 2D obtidas permitem com elevado detalhe morfológico e, num tempo de aquisição muito curto, a observação de todo o trato vocal e da posição e forma dos articuladores.

A Tabela 2 apresenta as medidas articulatórias realizadas nas imagens de cada sujeito do sexo feminino (nº 1 e 2) e masculino (nº 3 e 4) para cada som em estudo. Adicionalmente, com o intuito de remover o efeito da variabilidade nas dimensões anatómicas, são também apresentadas as médias de cada medida articulatória e respetivos coeficientes de variação. Assim, num estudo intersujeitos é possível observar quais as estratégias mais comumente adotadas na produção de cada som, nomeadamente a postura articulatória e a variação de cada medida.

Como se pode observar nos resultados da Tabela 2:

- Na produção do som /i/: existe uma grande variabilidade intersujeitos para as medidas “DLV” e “ORO”, sendo os coeficientes de variação iguais a 0,5 e 0,3, respetivamente, o que indica que estas se encontram

consideravelmente aumentadas comparativamente às restantes. O aumento destas duas medidas articulatórias origina uma maior abertura na parte posterior do trato vocal. As restantes medidas articulatórias encontram-se claramente diminuídas em todos os sujeitos. A variação relativa dos valores da medida “AL” é extremamente heterogénea e, por isso, uma característica extremamente variável intersujeitos;

- Na produção do som /a/: existe uma grande variabilidade para a medida articulatória “LAL”, apresentando um coeficiente de variação igual a 0,8, o que indica que esta se encontra extremamente aumentada comparativamente às restantes. A variabilidade é também moderadamente mais alta (30%) para as medidas articulatórias “AL” e “LDL”, evidenciando um aumento destas medidas comparativamente às outras medidas;
- Na produção do som /u/: existe uma variabilidade extremamente grande para a medida articulatória “LAL”, sendo o coeficiente de variação igual a 1,7, o que indica que esta se encontra extremamente aumentada comparativamente às restantes. A variabilidade nas restantes medidas é claramente inferior (à exceção da medida “ORO”), observando-se assim uma tendência oposta relativamente às médias dos valores obtidos. No que respeita à medida articulatória “ORO” é maior apenas para os sujeitos do sexo feminino; no entanto, a variação relativa intersujeitos é menor.

No que respeita ao estudo estático 3D, após segmentação das imagens e reconstrução de todo o trato vocal, determinaram-se os respetivos volumes associados a cada vogal e para cada um dos sujeitos nos contextos de voz normal e na simulação do padrão articulatório patológico de voz tensa (*cf.* Tabela 3).



Tabela 2: Resultados das medidas articulatórias realizadas nas imagens sagitais de RM: correlação intersujeitos e para cada som em estudo

SONS	SUJEITO	Medidas articulatórias (cm)											
		AL	Varição relativa (%)	CV	LAL	Varição relativa (%)	CV	LDL	Varição relativa (%)	CV	DLV	Varição relativa (%)	CV
/i/	1	1,28	2,9		0,76	-52,4		0,38	-52,5		2,54	137,4	
	2	0,76	1,3		0,29	-64,2		0,13	-75		2,29	48,4	
	3	0,29	-48,2	-5,2	0,52	-60	-0,2	0,24	-57,9	-0,2	0,91	200	0,3
	4	0,98	21		0,75	-42,2		0,27	-66,4		1,42	147,7	
	Média		-5,8			-54,7			-63,0			133,4	
/a/	1	2,2	76,9		2,28	42,8		1,6	100		0,19	-82,2	
	2	1,3	73,3		1,46	80,2		1,14	119,2		0,54	-65	
	3	0,84	50	0,3	1,77	36,2	0,8	0,88	54,4	0,3	0	-100	-0,2
	4	1,16	43,2		1,32	1,8		1,72	114,1		0,3	-47,7	
	Média		60,9			40,3			96,9			-73,7	
/u/	1	0,25	-79,9		1,75	9,6		0,42	-47,5		0,48	-55,1	
	2	0,19	-74,7		0,68	-16		0,29	-44,2		1,8	16,6	
	3	0,55	-1,8	-0,7	1,61	23,8	1,7	0,59	3,5	-0,7	0	-100	-0,9
	4	0,29	-64,2		1,82	40,4		0,42	-47,7		0	-100	
	Média		-55,2			14,5			-34,0			-59,6	
/l/-la/-lu/	Média de cada medida articulatória (cm)												
	SUJEITO	AL			LAL			LDL			DLV		
	1	1,2			1,6			0,8			1,1		
	2	0,8			0,8			0,5			1,5		
	3	0,6			1,3			0,3			0,3		
	4	0,8			1,3			0,6			0,6		

Abreviaturas: AL – abertura labial; LAL – localização do ápex da língua; LDL – localização do dorso da língua; DLV – distância da língua-véu do palato; ORO – abertura da orofaringe; CV – coeficiente de variação.

**Tabela 3:** Volumes do trato vocal obtidos a partir do estudo 3D estático por RM nos contextos de voz normal e voz tensa

Sujeito	Voz	Volumes do trato vocal (mm <sup>3</sup> )			Média /i/-/a/-/u/
		Som /i/	Som /a/	Som /u/	
1	normal	36864,0	48958,6	69554,0	51792,2
2		35303,0	28779,8	47518,9	37200,6
3		45080,0	41320,0	71331,1	52577,0
4	normal	49557,2	45275,1	76188,1	57006,8
	tensa	41093,3	41502,0	48080,6	
	diferença	-17,08%	-8,33%	-36,89%	
		Som /i/	Som /a/	Som /u/	
Variação relativa (%)		-28,8	-5,5	34,3	
		-5,1	-22,6	27,7	
		-14,3	-21,4	35,7	
		-13,1	-20,6	33,6	
Média		-15,3	-17,5	32,8	
Coeficiente de variação		-0,6	-0,5	0,1	

No contexto de voz falada intersujeitos é notória a variabilidade nos volumes estimados do trato vocal para cada som, destacando-se, no entanto, uma tendência comum para a vogal /u/, cujos volumes medidos foram consideravelmente maiores para todos os sujeitos por comparação com as outras vogais. Seguiram-se maiores volumes do trato vocal para a vogal /i/ e, por fim, a vogal /a/, à exceção do sujeito 1, em que se verificou o oposto.

Pela análise dos coeficientes de variação, pode concluir-se que o volume do trato vocal surge consistentemente aumentado, em cerca de duas vezes, na produção da vogal /u/, havendo uma maior variabilidade intersujeitos quando comparados os valores obtidos para as outras vogais.

Quanto ao padrão patológico de voz tensa, apenas para o sujeito 4 os volumes do trato vocal e principais cavidades de ressonância são claramente inferiores e mais expressivas para a vogal /u/; ou seja, condicionando a redução dos volumes comparativamente à voz falada desse sujeito.

No padrão patológico de voz tensa, as diferenças entre os volumes medidos durante a produção da vogal /a/ são menores; ou seja, existe uma redução do volume de ressonância por comparação à voz normal do sujeito.

No que respeita ao padrão patológico do sigmatismo obtiveram-se três aquisições volumétricas do sujeito 4 durante a produção da sílaba /sa/ nos contextos de voz normal e sigmatismo (interdental e lateral). A Figura 2 ilustra as diferenças observadas nos modelos 3D do trato vocal e respetivos volumes extraídos.

Como pode observar-se na Figura 2, no caso do sigmatismo interdental, o volume avaliado do trato vocal é

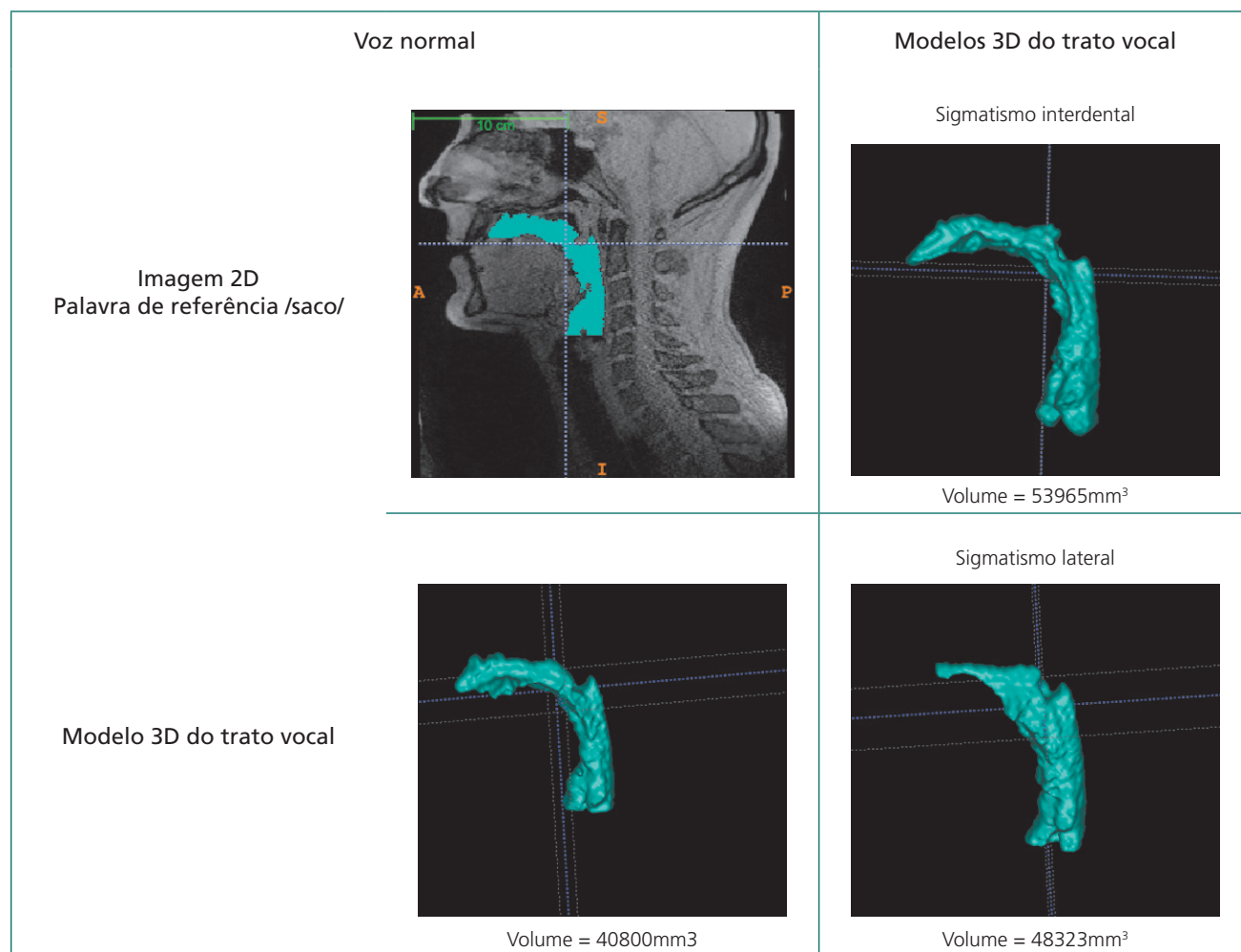
cerca de 32,2% superior comparativamente à voz normal do sujeito, usando a mesma palavra de referência. Já para o padrão patológico de sigmatismo lateral, a diferença é inferior (18,4%); no entanto, também revelou um aumento do volume de todo o trato vocal.

Por fim, num estudo com objetivo de avaliar qual o efeito da sustentação prolongada (superior a 30 segundos) dos sons e da posição corporal (decúbito dorsal) imposta aos sujeitos durante as aquisições de RM<sup>29</sup> observou-se que a conjugação destes dois efeitos implica a hiperarticulação (exagero) da fala. Segundo este autor, a sustentação prolongada dos sons exige um grande esforço por parte dos sujeitos, justificando-se, por isso, a utilização de tempos de aquisição muito mais curtos e, se possível, em contexto real. O tempo de aquisição apresentado no protocolo de RM que propomos retrata essa preocupação.

Consideramos assim que, apesar de se ter optado pela sustentação dos sons durante as aquisições de RM, ainda que num tempo de aquisição bastante mais curto, foi compensado pelo facto dos sujeitos serem profissionais da voz e apresentarem um domínio e treino vocal muito amplo, garantindo assim um correto posicionamento dos articuladores e não comprometendo a fiabilidade dos resultados.

## Conclusão

A utilização deste protocolo de RM permitiu a recolha de um conjunto de dados morfológicos qualitativos e quantitativos para a caracterização de três vogais do PE e ainda a simulação de dois padrões patológicos articulatorios.



**Figura 2:** Modelos 3D do trato vocal e respectivos volumes obtidos a partir do estudo 3D estático por RM num participante do sexo masculino, nos contextos de voz normal e patológica (simulada) de sigmatismo.

Da análise morfológica 2D do trato vocal foi possível identificar quais as estratégias mais comumente adotadas na produção de cada som, nomeadamente a postura articulatória e a variação de cada medida para cada um dos sujeitos. Na produção do som /i/ foi encontrada uma grande variabilidade intersujeitos para as medidas “DLV” e “ORO”, sendo os coeficientes de variação iguais a 0,5 e 0,3, respetivamente, o que indica que estas se encontram consideravelmente aumentadas comparativamente às restantes. Nesta vogal em particular constatou-se que a abertura labial é uma característica extremamente variável entre os sujeitos. A postura articulatória adotada pelos sujeitos na produção da vogal /a/ é a mais consistente no que respeita aos valores da variabilidade relativa das medidas articulatórias. Na vogal /u/ observou-se um aumento considerável da medida “LAL” comparativamente às restantes.

Relativamente à análise morfométrica 3D foram reconstruídos 17 modelos 3D do trato vocal a partir das imagens de RM. No contexto de voz falada intersujeitos foi notória a variabilidade nos volumes estimados do trato vocal para cada som e, em especial, o aumento do volume do trato vocal na perturbação articulatória de sigmatismo.

O desenvolvimento de técnicas para visualização dos dentes nas imagens de RM é também uma das linhas de investigação cada vez mais visível<sup>30-35</sup>, no sentido de aumentar o rigor das medidas volumétricas do trato vocal extraídas. Estas técnicas assentam, principalmente, na utilização de meios de contraste na cavidade oral<sup>33</sup> ou no desenvolvimento de moldes/dispositivos para identificação dos dentes<sup>31-32,35</sup>, a partir de sequências de pulso convencionais de RM (spin-eco, gradiente-eco e inversão-recuperação). Destaca-se, por fim, o desenvolvimento de uma sequência inovadora de RM obtida num sistema de 4.0 Tesla, denominada de *SWEEP Imaging with Fourier Transform* (SWIFT)<sup>30</sup> e que permite a observação 3D simultânea dos tecidos moles e duro (mineralizado) do dente.

No futuro perspetiva-se a utilização de novas técnicas para identificação dos dentes, abrindo novos horizontes de estudo para outras áreas, como a ortodontia e a odontologia.

Em suma, apesar das condicionantes que ainda se verificam, a imagem por RM é sem dúvida uma técnica promissora no estudo da fala, inócua, não-invasiva e que fornece informação fiável e com utilidade não só em engenharia, em medicina e em terapia da fala, mas também em diversas outras áreas.



## Agradecimentos

As imagens de ressonância magnética foram adquiridas na Unidade de RM do Centro Hospitalar de S. João do Porto, com a colaboração de toda a equipa técnica, a quem agradecemos.

A primeira autora agradece os apoios concedidos no âmbito da atribuição da bolsa de doutoramento pela Escola Superior de Tecnologia da Saúde do Porto e pelo Instituto Politécnico do Porto.

## Referências bibliográficas

- Rokkaku M, Hashimoto K, Imaizumi S, Niimi S, Kiritani S. Measurements of the three-dimensional shape of the vocal tract based on the magnetic resonance imaging technique. *Ann Bull Res Inst Logoped Phoniater*. 1986;20:47-54.
- Demolin D, Metens T, Soquet A. Three-dimensional measurement of the vocal tract by MRI. In *Proceedings of the 4<sup>th</sup> International Conference on Spoken Language Processing (ICSLP 96)*. Philadelphia, USA; 1996. p. 272-5.
- Badin P, Bailly G, Revéret L, Baciú M, Segebarth C, Savariaux C. Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images. *J Phon*. 2002;30(3):533-53.
- Narayanan S, Nayak K, Lee S, Sethy A, Byrd D. An approach to real-time magnetic resonance imaging for speech production. *J Acoust Soc Am*. 2004;115(4):1771-6.
- Story BH. Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. *J Acoust Soc Am*. 2008;123(1):327-35.
- Birkholz P, Kröger BJ. Vocal tract model adaptation using magnetic resonance imaging. In *Proceedings of the 7<sup>th</sup> International Seminar on Speech Production (ISSP'06)*. Ubatuba, Belo Horizonte, Brazil; 2006. p. 493-500.
- Takemoto H, Honda K, Masaki S, Shimada Y. Measurement of temporal changes in vocal tract area function during a continuous vowel sequence using a 3D Cine-MRI technique. In *Proceedings of the 6<sup>th</sup> International Seminar on Speech Production*. Sydney, Australia; 2003. p. 284-9.
- Engwall O. Combining MRI, EMA and EPG measurements in a three-dimensional tongue model. *Speech Commun*. 2003;41(2-3):303-29.
- Engwall O. A 3D tongue model based on MRI data. In *Proceedings of the Sixth International Conference on Spoken Language Processing*. Beijing, China; 2000. p. 1-4.
- Ramanarayana V, Goldstein L, Byrd D, Narayanan S. An MRI study of articulatory settings of L1 and L2 speakers of American English. In *Proceedings of the 9<sup>th</sup> International Seminar on Speech Production (ISSP'11)*. Montreal, Canada; 2011.
- Kröger BJ, Winkler R, Mooshammer C, Pompino-Marschall B. Estimation of vocal tract area function from magnetic resonance imaging: preliminary results. In *Proceedings of the 5<sup>th</sup> Seminar on Speech Production*. München, Germany; 2000. p. 333-6.
- Doel K Van Den, Vogt F, English R. Towards articulatory speech synthesis with a dynamic 3D finite element tongue model. In *Proceedings of the 7<sup>th</sup> International Seminar on Speech Production (ISSP'06)*. Ubatuba, Belo Horizonte, Brazil; 2006. p. 59-66.
- Bresch E, Adams J, Pouzet A, Lee S, Byrd D, Narayanan S. Semi-automatic processing of real-time MR image sequences for speech production studies. In *Proceedings of the 7<sup>th</sup> International Seminar on Speech Production (ISSP'06)*. Ubatuba, Belo Horizonte, Brazil; 2006. p. 427-34.
- Narayanan SS, Alwan AA, Haker K. An articulatory study of fricative consonants using magnetic resonance imaging. *J Acoust Soc Am*. 1995;98(3):1325-47.
- Fricke BL, Abbott MB, Donnelly LF, Dardzinski BJ, Poe SA, Karl M, et al. Upper airway volume segmentation analysis using cine MRI findings in children with tracheostomy tubes. *Korean J Radiol*. 2007;8(6):506-11.
- Mády K, Sader R, Zimmermann A, Hoole P, Beer A, Zeilhofer HF, et al. Assessment of consonant articulation in glossectomy speech by dynamic MRI. In *Proceedings of the 7<sup>th</sup> International Conference on Spoken Language Processing (ICSLP)*. Denver; 2002. p. 3-6.
- Mády K, Sader R, Zimmermann A, Hoole P, Beer A, Zeilhofer H, et al. Use of real-time MRI in assessment of consonant articulation before and after tongue surgery and tongue reconstruction. In *Proceedings of the 4<sup>th</sup> International Speech Motor Conference*. Nijmegen, The Netherlands; 2001. p. 142-5.
- Masaki S, Nota Y, Takano S, Takemoto H, Kitamura T, Honda K. Integrated magnetic resonance imaging methods for speech science and technology. *J Acoust Soc Am*. 2008;123(5):375-9.
- Rua SM, Freitas DR. Morphological dynamic imaging of human vocal tract. In *Proceedings of the Computational Modelling of Objects Represented in Images: fundamentals, methods and applications (ComplIMAGE)*. Coimbra, Portugal; 2006. p. 381-6.
- Martins P, Carbone I, Pinto A, Silva A, Teixeira A. European Portuguese MRI based speech production studies. *Speech Commun*. 2008;50(11-12):925-52.
- Serrurier A, Badin P. Towards a 3D articulatory model of velum based on MRI and CT images. *ZAS Pap Linguist*. 2005;40(1):195-211.
- Ventura SR, Freitas DR, Tavares JM. Application of MRI and biomedical engineering in speech production study. *Comput Methods Biomech Biomed Engin*. 2009;12(6):671-81.
- Martins P, Oliveira C, Silva A, Inform T. Articulatory characteristics of European Portuguese laterals: a 2D & 3D MRI study. In *VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop*. Vigo, Spain; 2010. p. 33-6.
- Ventura SR, Freitas DR, Tavares JM. Toward dynamic magnetic resonance imaging of the vocal tract during speech production. *J Voice*. 2011;25(4):511-8.
- Vasconcelos MJ, Ventura SR, Freitas DR, Tavares JM. Towards the automatic study of the vocal tract from magnetic resonance images. *J Voice*. 2011;25(6):732-42.
- Ventura SR, Freitas DR, Ramos IM, Tavares JM. Requisitos e condicionantes da imagem por ressonância magnética no estudo da fala humana. In *Congresso de Métodos Numéricos em Engenharia (CMNE)*. Coimbra, Portugal; 2011. p. 1-12.
- Vasconcelos MJ, Ventura SR, Freitas DR, Tavares JM. Using statistical deformable models to reconstruct vocal tract shape from magnetic resonance images. In *Proceedings of*

- the Institution of Mechanical Engineers, Part H. *J Engin Med.* 2010;224(10):1153-63.
28. Ventura SR, Freitas DR, Ramos IM, Tavares JM. Morphologic differences in the vocal tract resonance cavities of voice professionals: an MRI-based study. *J Voice.* 2013;27(2):132-40.
  29. Engwall O. Assessing MRI measurements: effects of sustentation, gravitation and coarticulation. In Harrington J, Tabain M, editors. *Speech production: models, phonetic processes and techniques.* New York: Psychology Press; 2006. p. 301-14. ISBN 9781841694375
  30. Idiyatullin D, Corum C, Moeller S, Prasad HS, Garwood M, Nixdorf DR. Dental magnetic resonance imaging: making the invisible visible. *J Endod.* 2011;37(6):745-52.
  31. Kitamura T, Nishimoto H, Fujimoto I, Shimada Y. Dental imaging using magnetic resonance visible mouthpiece for measurement of vocal tract shape and dimensions. *Acoust Sci Technol.* 2011;32(5):224-7.
  32. Ng IW, Ono T, Inoue-Arai MS, Honda E, Kurabayashi T, Moriyama K. Application of MRI movie for observation of articulatory movement during fricative /s/ and a plosive /t/. *Angle Orthod.* 2011;81(2):237-44.
  33. Olt S, Jakob PM. Contrast-enhanced dental MRI for visualization of the teeth and jaw. *Magn Reson Med.* 2004;52(1):174-6.
  34. Tutton LM, Goddard PR. MRI of the teeth. *Br J Radiol.* 2002;75(894):552-62.
  35. Ventura SR, Freitas DR, Ramos IM, Tavares JM. Three-dimensional visualization of teeth by magnetic resonance imaging during speech. In *Proceedings of the II International Conference on Biodental Engineering.* Porto, Portugal; 2012. p. 13-7.

Artigo recebido em 15.05.2013 e aprovado em 08.05.2014